

Channel Coding in Random Access Communication over Compound Channels

Zheng Wang, Jie Luo

Abstract

Due to the short and bursty incoming messages, channel access activities in a wireless random access system are often fractional. The lack of frequent data support consequently makes it difficult for the receiver to estimate and track the time varying channel states with high precision. This paper investigates random multiple access communication over a compound wireless channel where channel realization is known neither at the transmitters nor at the receiver. An achievable rate and error probability tradeoff bound is derived under the non-asymptotic assumption of a finite codeword length. The results are then extended to the random multiple access system where the receiver is only interested in decoding messages from a user subset.

Index Terms

channel coding, compound channel, finite codeword length, random access

I. INTRODUCTION

In random multiple access communication, users (transmitters) determine their communication rates individually, without sharing the rate information either among each other or with the receiver [1]. With the absence of rate coordination among users, reliable message recovery is not always possible [2]. The receiver in this case decodes the transmitted messages if a pre-determined error probability requirement can be satisfied, or reports a collision otherwise [1].

Information theoretic channel coding in time-slotted random multiple access communication over a discrete-time memoryless channel was recently investigated in [1][3]. Assume that channel coding is applied only within each time slot (or packet). It was shown in [1] that the fundamental performance limitation of the system can be characterized using an achievable rate region in the following sense. Asymptotically as the codeword length (or time slot length) is taken to infinity, the receiver is able to

The authors are with the Electrical and Computer Engineering Department, Colorado State University, Fort Collins, CO 80523. E-mail: {zhwang, rocky}@engr.colostate.edu.

This work was supported by the National Science Foundation under Grants CCF-1016985 and CNS-1116134. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

recover the messages reliably if the communication rate vector (which includes the rates of all users) happens to be inside the rate region, and to reliably report a collision if the rate vector happens to be outside the region [1]. The achievable rate region was shown to coincide with the Shannon information rate region without a convex hull operation [1]. In [3], the asymptotic coding result was further strengthened to a rate and error probability tradeoff bound under the assumption of a finite codeword length. A bound on the achievable error exponent was obtained consequently [3].

Both [1] and [3] assumed that the channel state information is known at the receiver. Unfortunately, since random access communication deals with bursty short messages, transmission activities of a user are often fractional. Without frequent data support, accurate real-time channel estimation and tracking become difficult at the receiver. Understanding the system performance limitation without channel state information therefore becomes essential [4]. In this paper, we illustrate how coding theorems developed in [1][3] can be extended to random multiple access communication over a compound discrete-time memoryless channel [5][6], consisting of a family (set) of channels over which the communication could take place. Both the transmitters and the receiver know about the compound channel set, but not the actual channel realization. As in [1][3], we assume that time is partitioned into slots of equal length, and we focus on channel coding within one time slot. We define the communication rate of a user as the normalized number of information nats encoded in a packet (or a time slot).

The compound channel communication problem investigated in this paper is different from a conventional one in the following two key aspects. First, in a conventional system, information rates are jointly determined by the transmitters and the receiver [7], while communication rates in a random access system are determined distributively and the rate information is unknown at the receiver [1]. Second, in a conventional system, in order to achieve reliable communication, the transmitted rate vector should be supported by all channel realizations in the compound set [8][4]. In random access communication, however, even though the receiver needs to guarantee the reliability of its decoding output, the receiver also has the additional choice of reporting a collision to avoid confusing the upper layer networking [9]. This therefore allows the transmitted rate vector to be supported only by a subset of channel realizations. If the actual channel realization belongs to this subset, the receiver should decode the messages. Otherwise, the receiver should report a collision. Clearly, the decoding and collision report decisions made at the receiver are affected jointly by the communication rates of the users and the actual channel realization.

To address these key aspects in the system model, we assume that the receiver chooses an “operation region”, which is a set of rate vector and channel realization pairs. If the transmitted rate vector and channel realization pair is within the operation region, the receiver *intends* to decode the messages, otherwise the receiver *intends* to report a collision (or outage). We define the decoding error probability and the collision

miss detection probability similarly to [3], and define the system error probability as the maximum of the two. An upper bound on the achievable system error probability is derived under the assumption of a finite codeword length. We then show how the compound channel results help in obtaining error performance bounds for the random multiple access system where the receiver is only interested in recovering messages from a user *subset* [1]. This is based on the observation that, conditioned on the receiver not decoding messages for the rest of the users, the impact of their communication activities on the user subset of interest is equivalent to that of a compound channel.

II. MULTIPLE RANDOM ACCESS COMMUNICATION OVER A COMPOUND CHANNEL

Consider a K -user time-slotted random access system over a compound discrete-time memoryless channel. Time is slotted with each slot equaling N symbol durations, which is also the length of a packet or a codeword. We assume that channel coding is only applied within each time slot or packet. The compound channel consists of a family of discrete-time memoryless channels, characterized by a set of conditional probabilities $\{P_{Y|X_1, \dots, X_K}^{(1)}, \dots, P_{Y|X_1, \dots, X_K}^{(H)}\}$ with cardinality H , where, for $k \in \{1, \dots, K\}$, $X_k \in \mathcal{X}$ is the channel input symbol of user k with \mathcal{X} being the finite input alphabet, and $Y \in \mathcal{Y}$ is the channel output symbol with \mathcal{Y} being the finite output alphabet. In each time slot, a channel realization is randomly generated from this set and remains static throughout the slot duration. We assume that all users and the receiver know the compound channel set, but not the actual channel realization. For the time being, we will assume that $H < \infty$. The case when the compound channel set contains an infinite number of channels will be discussed at the end of this section.

Assume that at the beginning of a time slot, according to the message availability and the MAC layer protocol, each user, say user k ($k \in \{1, \dots, K\}$) chooses an arbitrary communication rate $r_k \in \{r_{k1}, \dots, r_{kM}\}$ in nats per symbol, where $\{r_{k1}, \dots, r_{kM}\}$ is a pre-determined finite rate set of user k with cardinality M . Neither the other users nor the receiver knows the actual rate realization for each transmission, although they are shared with the rate set information. The user then encodes $\lfloor Nr_k \rfloor$ number of data nats, denoted by a message w_k , into a packet (codeword) with N symbols, using a random coding scheme specified as in [1][10] and also in the following. For all $k \in \{1, \dots, K\}$, we assume that user k is equipped with a codebook library $\mathcal{L}_k = \{\mathcal{C}_{k\theta_k} : \theta_k \in \Theta_k\}$ in which codebooks are indexed by a set Θ_k . Each codebook has M classes of codewords. The i^{th} ($i \in \{1, \dots, M\}$) codeword class has $\lfloor e^{Nr_{ki}} \rfloor$ codewords with the same length of N symbols. In contrast to a conventional coding scheme, here each codeword in the codebook corresponds to a message and rate pair (w_k, r_k) [1][3]. Let $\mathcal{C}_{k\theta_k}(w_k, r_k)_j$ be the j^{th} symbol of the codeword corresponding to message and rate pair (w_k, r_k) in codebook $\mathcal{C}_{k\theta_k}$. User k first selects the codebook by generating θ_k according to a distribution ϑ_k such that the random variables

$X_{(w_k, r_k), j} : \theta_k \rightarrow \mathcal{C}_{k\theta_k}(w_k, r_k)_j$ are i.i.d. according to an input distribution $P_{X|r_k}$ ¹. The codebook $\mathcal{C}_{k\theta_k}$ is then used to map (w_k, r_k) into a codeword, denoted by $\mathbf{x}_{(w_k, r_k)}$. After encoding, the codewords of all users are sent to the receiver over the compound channel.

To simplify the notation, we use bold font variable to denote the corresponding variables of all users. For example, \mathbf{w} and \mathbf{r} denote the messages and communication rates of all users. $\mathbf{P}_{\mathbf{X}|\mathbf{r}}$ denote the input distributions of all users, etc. Given a vector variable, say \mathbf{r} , we use r_i to denote its element corresponding to user i . Let $\mathcal{S} \subset \{1, \dots, K\}$ be a user subset, and $\bar{\mathcal{S}}$ be its complement. We use $\mathbf{r}_{\mathcal{S}}$ to denote the vector that is extracted from \mathbf{r} with only elements corresponding to users in \mathcal{S} . By using the vector notation of the channel input symbols, the compound channel set is also denoted by $\{P_{Y|\mathbf{X}}^{(1)}, \dots, P_{Y|\mathbf{X}}^{(H)}\}$.

We assume that the receiver is shared with the random codebook generation algorithms and hence knows the randomly generated codebooks of all users. Before packet transmission, the receiver pre-determines an “operation region” $\mathcal{R} = \{(\mathbf{r}, P_{Y|\mathbf{X}})\}$, which is a set of rate vector and channel realization pair, where each entry of \mathbf{r} is chosen from the corresponding rate set, i.e., $r_k \in \{r_{k1}, \dots, r_{kM}\}$ ($k \in \{1, \dots, K\}$), and $P_{Y|\mathbf{X}} \in \{P_{Y|\mathbf{X}}^{(1)}, \dots, P_{Y|\mathbf{X}}^{(H)}\}$. Let $(\mathbf{r}, P_{Y|\mathbf{X}})$ be the actual realization of the transmitted rate vector and channel pair. We assume that the receiver *intends* to decode all messages if $(\mathbf{r}, P_{Y|\mathbf{X}}) \in \mathcal{R}$, and *intends* to report a collision if $(\mathbf{r}, P_{Y|\mathbf{X}}) \notin \mathcal{R}$. Note that the actual rate and channel realization $(\mathbf{r}, P_{Y|\mathbf{X}})$ is unknown at the receiver. Therefore the receiver needs to make decisions whether to decode messages or to report a collision only based on the received channel symbols. More specifically, in each time slot, upon receiving the channel output symbols \mathbf{y} , the receiver estimates the rate and channel pair, denoted by $(\hat{\mathbf{r}}, \hat{P}_{Y|\mathbf{X}})$, for all users. The receiver outputs the corresponding estimated message and rate vector pair $(\hat{\mathbf{w}}, \hat{\mathbf{r}})$ if $(\hat{\mathbf{r}}, \hat{P}_{Y|\mathbf{X}}) \in \mathcal{R}$ and a pre-determined decoding error probability requirement is satisfied. Otherwise, the receiver reports a collision. Also note that, whether the receiver should recover the messages or report a collision not only depends on the rates, but also depends on the channel realization. In other words, for the same transmission rate vector, the receiver may be designed to take different actions for different channel realizations. This is opposed to the conventional compound channel communication scenario where, if a rate is supported by the system, the receiver should always decode the messages irrespective of the channel realization.

Given the operation region \mathcal{R} , and conditioned on that (\mathbf{w}, \mathbf{r}) is transmitted over channel $P_{Y|\mathbf{X}}$, we define the following three error probabilities. The decoding error probability, for $(\mathbf{w}, \mathbf{r}, P_{Y|\mathbf{X}})$ with $(\mathbf{r}, P_{Y|\mathbf{X}}) \in \mathcal{R}$, is defined as

$$P_{e(\mathbf{w}, \mathbf{r}, P_{Y|\mathbf{X}})} = Pr \left\{ (\hat{\mathbf{w}}, \hat{\mathbf{r}}) \neq (\mathbf{w}, \mathbf{r}) | (\mathbf{w}, \mathbf{r}, P_{Y|\mathbf{X}}) \right\}, \quad \forall (\mathbf{w}, \mathbf{r}, P_{Y|\mathbf{X}}), (\mathbf{r}, P_{Y|\mathbf{X}}) \in \mathcal{R}. \quad (1)$$

¹The input distribution is assumed to be a function of the communication rate. In other words, different communication rates may correspond to different input distributions.

The collision miss detection probability, for $(\mathbf{w}, \mathbf{r}, P_{Y|\mathbf{X}})$ with $(\mathbf{r}, P_{Y|\mathbf{X}}) \notin \mathcal{R}$, is defined as

$$\bar{P}_{c(\mathbf{w}, \mathbf{r}, P_{Y|\mathbf{X}})} = 1 - \Pr \left\{ \text{"collision"} | (\mathbf{w}, \mathbf{r}, P_{Y|\mathbf{X}}) \right\} - \Pr \left\{ (\hat{\mathbf{w}}, \hat{\mathbf{r}}) = (\mathbf{w}, \mathbf{r}) | (\mathbf{w}, \mathbf{r}, P_{Y|\mathbf{X}}) \right\},$$

$$\forall (\mathbf{w}, \mathbf{r}, P_{Y|\mathbf{X}}), (\mathbf{r}, P_{Y|\mathbf{X}}) \notin \mathcal{R}. \quad (2)$$

Note that in (2), when $(\mathbf{r}, P_{Y|\mathbf{X}}) \notin \mathcal{R}$, we have excluded the correct message and rate pair estimation from the collision miss detection event.

Let $\mathcal{S} \subset \{1, \dots, K\}$ be an arbitrary user subset. Assume that $\sum_{k \notin \mathcal{S}} r_k \leq I_{(\mathbf{r}, P_{Y|\mathbf{X}})}(\mathbf{X}_{\bar{\mathcal{S}}}; \mathbf{Y} | \mathbf{X}_{\mathcal{S}})$ for all $(\mathbf{r}, P_{Y|\mathbf{X}}) \in \mathcal{R}$, where $\mathbf{X}_{\mathcal{S}}$ denotes the channel input symbols of users in set \mathcal{S} , and $\mathbf{X}_{\bar{\mathcal{S}}}$ denotes the channel input symbols of users not in set \mathcal{S} . $I_{(\mathbf{r}, P_{Y|\mathbf{X}})}$ is the mutual information function computed using input distribution corresponding to rate vector \mathbf{r} (i.e., $\mathbf{P}_{\mathbf{X}|\mathbf{r}}$) and channel $P_{Y|\mathbf{X}}$. We define the system error probability P_{es} as

$$P_{es} = \max \left\{ \max_{(\mathbf{w}, \mathbf{r}, P_{Y|\mathbf{X}}), (\mathbf{r}, P_{Y|\mathbf{X}}) \in \mathcal{R}} P_{e(\mathbf{w}, \mathbf{r}, P_{Y|\mathbf{X}})}, \max_{(\mathbf{w}, \mathbf{r}, P_{Y|\mathbf{X}}), (\mathbf{r}, P_{Y|\mathbf{X}}) \notin \mathcal{R}} \bar{P}_{c(\mathbf{w}, \mathbf{r}, P_{Y|\mathbf{X}})} \right\}. \quad (3)$$

The following theorem gives an upper bound on the achievable system error probability P_{es} .

Theorem 1: Consider K -user multiple random access communication over a compound discrete-time memoryless channel $\{P_{Y|\mathbf{X}}^{(1)}, \dots, P_{Y|\mathbf{X}}^{(H)}\}$, where $H < \infty$ is a positive integer. Let $\mathbf{P}_{\mathbf{X}|\mathbf{r}}$ be the input distribution for all users and all rates. Let \mathcal{R} be the operation region. Assume finite codeword length N . There exists a decoding algorithm, whose system error probability P_{es} is upper bounded by,

$$P_{es} \leq \max \left\{ \max_{(\mathbf{r}, P_{Y|\mathbf{X}}) \in \mathcal{R}} \sum_{\mathcal{S} \subset \{1, \dots, K\}} \left[\sum_{(\tilde{\mathbf{r}}, \tilde{P}_{Y|\mathbf{X}}) \in \mathcal{R}, \tilde{\mathbf{r}}_S = \mathbf{r}_S} \exp\{-N E_m(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}}, P_{Y|\mathbf{X}}, \tilde{P}_{Y|\mathbf{X}})\} \right. \right. \\ \left. \left. + \max_{(\mathbf{r}', P'_{Y|\mathbf{X}}) \notin \mathcal{R}, \mathbf{r}'_S = \mathbf{r}_S} \exp\{-N E_i(\mathcal{S}, \mathbf{r}, \mathbf{r}', P_{Y|\mathbf{X}}, P'_{Y|\mathbf{X}})\} \right] \right. \\ \left. \max_{(\tilde{\mathbf{r}}, \tilde{P}_{Y|\mathbf{X}}) \notin \mathcal{R}} \sum_{\mathcal{S} \subset \{1, \dots, K\}} \sum_{(\mathbf{r}, P_{Y|\mathbf{X}}) \in \mathcal{R}, \mathbf{r}_S = \tilde{\mathbf{r}}_S} \max_{(\mathbf{r}', P'_{Y|\mathbf{X}}) \notin \mathcal{R}, \mathbf{r}'_S = \tilde{\mathbf{r}}_S} \exp\{-N E_i(\mathcal{S}, \mathbf{r}, \mathbf{r}', P_{Y|\mathbf{X}}, P'_{Y|\mathbf{X}})\} \right\}, \quad (4)$$

where $E_m(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}}, P_{Y|\mathbf{X}}, \tilde{P}_{Y|\mathbf{X}})$ and $E_i(\mathcal{S}, \mathbf{r}, \mathbf{r}', P_{Y|\mathbf{X}}, P'_{Y|\mathbf{X}})$ are given by

$$E_m(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}}, P_{Y|\mathbf{X}}, \tilde{P}_{Y|\mathbf{X}}) = \max_{0 < \rho \leq 1} -\rho \sum_{k \notin \mathcal{S}} \tilde{r}_k + \max_{0 < s \leq 1} -\log \sum_Y \sum_{\mathbf{X}_{\mathcal{S}}} \prod_{k \in \mathcal{S}} P_{X|r_k}(X_k) \\ \times \left(\sum_{\mathbf{X}_{\bar{\mathcal{S}}}} \prod_{k \notin \mathcal{S}} P_{X|r_k}(X_k) P_{Y|\mathbf{X}}(Y|\mathbf{X})^{1-s} \right) \left(\sum_{\mathbf{X}_{\bar{\mathcal{S}}}} \prod_{k \notin \mathcal{S}} P_{X|\tilde{r}_k}(X_k) \tilde{P}_{Y|\mathbf{X}}(Y|\mathbf{X})^{\frac{s}{\rho}} \right)^{\rho}, \\ E_i(\mathcal{S}, \mathbf{r}, \mathbf{r}', P_{Y|\mathbf{X}}, P'_{Y|\mathbf{X}}) = \max_{0 < \rho \leq 1} -\rho \sum_{k \notin \mathcal{S}} r_k + \max_{0 < s \leq 1-\rho} -\log \sum_Y \sum_{\mathbf{X}_{\mathcal{S}}} \prod_{k \in \mathcal{S}} P_{X|r_k}(X_k) \\ \times \left(\sum_{\mathbf{X}_{\bar{\mathcal{S}}}} \prod_{k \notin \mathcal{S}} P_{X|r_k}(X_k) P_{Y|\mathbf{X}}(Y|\mathbf{X})^{\frac{s}{s+\rho}} \right)^{s+\rho} \left(\sum_{\mathbf{X}_{\bar{\mathcal{S}}}} \prod_{k \notin \mathcal{S}} P_{X|r'_k}(X_k) P'_{Y|\mathbf{X}}(Y|\mathbf{X}) \right)^{1-s}. \quad (5)$$

The proof of Theorem 1 is given in Appendix A.

When the compound channel is randomly generated at the beginning but remains static afterwards, one can take codeword length to infinity to obtain the system error exponent as $E_s = \lim_{N \rightarrow \infty} -\frac{1}{N} \log P_{es}$. The

following lower bound on the achievable system error exponent E_s can be easily derived from Theorem 1.

Corollary 1: The system error exponent of a K -user multiple random access system over compound discrete-time memoryless channels given in Theorem 1 is lower-bounded by

$$E_s \geq \min \left\{ \min_{S \subset \{1, \dots, K\}} \min_{(\mathbf{r}, P_{Y|\mathbf{X}}), (\tilde{\mathbf{r}}, \tilde{P}_{Y|\mathbf{X}}) \in \mathcal{R}} E_m(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}}, P_{Y|\mathbf{X}}, \tilde{P}_{Y|\mathbf{X}}), \right. \\ \left. \min_{S \subset \{1, \dots, K\}} \min_{(\mathbf{r}, P_{Y|\mathbf{X}}) \in \mathcal{R}, (\tilde{\mathbf{r}}, \tilde{P}_{Y|\mathbf{X}}) \notin \mathcal{R}} E_i(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}}, P_{Y|\mathbf{X}}, \tilde{P}_{Y|\mathbf{X}}) \right\}, \quad (6)$$

where $E_m(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}}, P_{Y|\mathbf{X}}, \tilde{P}_{Y|\mathbf{X}})$ and $E_i(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}}, P_{Y|\mathbf{X}}, \tilde{P}_{Y|\mathbf{X}})$ are given in (5).

Compared with the error exponent derived in [3, Corollary 2], we can see that, even though the channel stays static forever, the system still needs to pay a penalty in error exponent performance for not knowing the channel at the receiver².

In both Theorem 1 and Corollary 1, we have assumed that there are only a finite number of channels in the compound set. Next, we will extend the result to the case when the cardinality of the compound channel set can be infinity.

We first assume that the channels in the compound set can be partitioned into H classes, denoted by $\{\mathcal{F}^{(1)}, \dots, \mathcal{F}^{(H)}\}$, where $H < \infty$ is a positive integer. For example, if the compound channel set contains fading channels with continuous channel gains, one could quantize the channel gains and define the set of channels with the same quantization outcome as one channel class. We next assume that the receiver should choose an operation region \mathcal{R} to satisfy the following constraint for any rate vector \mathbf{r} and channel class $\mathcal{F} \in \{\mathcal{F}^{(1)}, \dots, \mathcal{F}^{(H)}\}$.

$$\text{C1: For any } (\mathbf{r}, \mathcal{F}), \text{ either } (\mathbf{r}, P_{Y|\mathbf{X}}) \in \mathcal{R} \ \forall P_{Y|\mathbf{X}} \in \mathcal{F}, \text{ or } (\mathbf{r}, P_{Y|\mathbf{X}}) \notin \mathcal{R} \ \forall P_{Y|\mathbf{X}} \in \mathcal{F}. \quad (7)$$

We say $(\mathbf{r}, \mathcal{F}) \in \mathcal{R}$ if $(\mathbf{r}, P_{Y|\mathbf{X}}) \in \mathcal{R}$ for all $P_{Y|\mathbf{X}} \in \mathcal{F}$, and we say $(\mathbf{r}, \mathcal{F}) \notin \mathcal{R}$ otherwise.

For each channel class \mathcal{F} and for each channel output symbol Y and input symbol vector \mathbf{X} , we define the following upper and lower bounds on the conditional probability values yielded by channels in \mathcal{F} , denoted by $P_{\max}^{\mathcal{F}}(Y|\mathbf{X})$ and $P_{\min}^{\mathcal{F}}(Y|\mathbf{X})$,

$$P_{\max}^{\mathcal{F}}(Y|\mathbf{X}) = \max_{P_{Y|\mathbf{X}} \in \mathcal{F}} P_{Y|\mathbf{X}}(Y|\mathbf{X}), \quad P_{\min}^{\mathcal{F}}(Y|\mathbf{X}) = \min_{P_{Y|\mathbf{X}} \in \mathcal{F}} P_{Y|\mathbf{X}}(Y|\mathbf{X}). \quad (8)$$

The following theorem gives an upper bound on the achievable system error probability.

Theorem 2: Consider a K -user multiple random access communication system over a compound discrete-time memoryless channel. Assume that the compound set is partitioned into H classes, denoted

²We assume that such a conclusion should be well known for the conventional compound channel communication. However, we are not able to find a reference that made such a clear statement.

by $\{\mathcal{F}^{(1)}, \dots, \mathcal{F}^{(H)}\}$, where H is a finite positive integer. Assume that the operation region \mathcal{R} satisfies constraint C1 given in (7). The system error probability P_{es} is upper bounded as follows.

$$P_{es} \leq \max \left\{ \max_{(\mathbf{r}, \mathcal{F}) \in \mathcal{R}} \sum_{\mathcal{S} \subset \{1, \dots, K\}} \left[\max_{(\mathbf{r}', \mathcal{F}') \notin \mathcal{R}, \mathbf{r}'_{\mathcal{S}} = \mathbf{r}_{\mathcal{S}}} \exp \{-NE_i(\mathcal{S}, \mathbf{r}, \mathbf{r}', \mathcal{F}, \mathcal{F}')\} \right. \right. \\ \left. \left. + \sum_{(\tilde{\mathbf{r}}, \tilde{\mathcal{F}}) \in \mathcal{R}, \tilde{\mathbf{r}}_{\mathcal{S}} = \mathbf{r}_{\mathcal{S}}} \exp \{-NE_m(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}}, \mathcal{F}, \tilde{\mathcal{F}})\} \right] \right. \\ \left. \max_{(\tilde{\mathbf{r}}, \tilde{\mathcal{F}}) \notin \mathcal{R}} \sum_{\mathcal{S} \subset \{1, \dots, K\}} \left[\sum_{(\mathbf{r}, \mathcal{F}) \in \mathcal{R}, \mathbf{r}_{\mathcal{S}} = \tilde{\mathbf{r}}_{\mathcal{S}}} \max_{(\mathbf{r}', \mathcal{F}') \notin \mathcal{R}, \mathbf{r}'_{\mathcal{S}} = \mathbf{r}_{\mathcal{S}}} \exp \{-NE_i(\mathcal{S}, \mathbf{r}, \mathbf{r}', \mathcal{F}, \mathcal{F}')\} \right] \right\}. \quad (9)$$

where $E_m(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}}, \mathcal{F}, \tilde{\mathcal{F}})$ and $E_i(\mathcal{S}, \mathbf{r}, \mathbf{r}', \mathcal{F}, \mathcal{F}')$ are given by

$$E_m(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}}, \mathcal{F}, \tilde{\mathcal{F}}) = \max_{0 < \rho \leq 1} -\rho \sum_{k \notin \mathcal{S}} \tilde{r}_k + \max_{0 < s \leq 1} -\log \sum_Y \sum_{\mathbf{X}_{\mathcal{S}}} \prod_{k \in \mathcal{S}} P_{X|r_k}(X_k) \\ \times \left(\sum_{\mathbf{X}_{\bar{\mathcal{S}}}} \prod_{k \notin \mathcal{S}} P_{X|r_k}(X_k) P_{\max}^{\mathcal{F}}(Y|\mathbf{X}) P_{\min}^{\mathcal{F}}(Y|\mathbf{X})^{-s} \right) \left(\sum_{\mathbf{X}_{\bar{\mathcal{S}}}} \prod_{k \notin \mathcal{S}} P_{X|\tilde{r}_k}(X_k) P_{\max}^{\tilde{\mathcal{F}}}(Y|\mathbf{X})^{\frac{s}{\rho}} \right)^{\rho}. \\ E_i(\mathcal{S}, \mathbf{r}, \mathbf{r}', \mathcal{F}, \mathcal{F}') = \max_{0 < \rho \leq 1} -\rho \sum_{k \notin \mathcal{S}} r_k + \max_{0 < s \leq 1-\rho} -\log \sum_Y \sum_{\mathbf{X}_{\mathcal{S}}} \prod_{k \in \mathcal{S}} P_{X|r_k}(X_k) \\ \times \left(\sum_{\mathbf{X}_{\bar{\mathcal{S}}}} \prod_{k \notin \mathcal{S}} P_{X|r_k}(X_k) P_{\max}^{\mathcal{F}}(Y|\mathbf{X}) P_{\min}^{\mathcal{F}}(Y|\mathbf{X})^{\frac{-\rho}{s+\rho}} \right)^{s+\rho} \left(\sum_{\mathbf{X}_{\bar{\mathcal{S}}}} \prod_{k \notin \mathcal{S}} P_{X|r'_k}(X_k) P_{\max}^{\mathcal{F}'}(Y|\mathbf{X}) \right)^{1-s}. \quad (10)$$

■

The proof of Theorem 2 is given in Appendix B. As shown in the proof that, in order to make decoding and collision report decisions, the receiver only needs to search over the finite number of channel classes using statistics $P_{\max}^{\mathcal{F}}$ and $P_{\min}^{\mathcal{F}}$ defined in (8), as opposed to searching among all possible channels.

III. INDIVIDUAL USER DECODING IN RANDOM MULTIPLE ACCESS COMMUNICATION

In Section II, we have assumed that the receiver either decodes messages or reports collisions for *all* users in the system. In practical applications, even though many users compete for the wireless channel, it is common that the receiver may not be interested in recovering messages for all of them. In this section, we show that the results obtained in Section II can help to derive error probability bounds in a random multiple access system where the receiver is only interested in recovering the messages from a user subset. However, to simplify the notations, we will only consider a special case when the communication channel is known at the receiver, and when the receiver is only interested in decoding for a single user. Generalizing the results to decoding for multiple users over a compound channel is straightforward.

Let the discrete-time memoryless channel be characterized by $P_{Y|X}$, which is known at the receiver. In each time slot, each user chooses a communication rate and encodes its message using the random

coding scheme described in Section II. The rate information is shared neither among the users nor with the receiver. We assume that the receiver is only interested in recovering the message for user $k \in \{1, \dots, K\}$. We assume that the receiver chooses an operation region \mathcal{R} , such that if the transmitted rate vector \mathbf{r} satisfies $\mathbf{r} \in \mathcal{R}$, the receiver intends to decode for user k , and if $\mathbf{r} \notin \mathcal{R}$, the receiver intends to report a collision for user k . It is important to note that, first, whether the receiver will be able to decode the message of user k , not only depends on the rate of user k , but also depends on the rate of other users. Therefore, the operation rate region \mathcal{R} should still be defined as a set of rate vector \mathbf{r} , as opposed to the rate of user k . Second, even though the receiver only cares about the message of user k , the receiver still has the option of decoding the messages for some other users if this helps to improve the communication performance of user k . This implies that, based upon the received symbols, the receiver will essentially need to make a decision on which subset of the messages should be decoded.

Due to the above understandings, we first define an elementary decoder, called the “ $(\mathcal{D}, \mathcal{R}_{\mathcal{D}})$ -decoder”. Given a user subset $\mathcal{D} \subseteq \{1, \dots, K\}$ and an operation rate region $\mathcal{R}_{\mathcal{D}}$, the “ $(\mathcal{D}, \mathcal{R}_{\mathcal{D}})$ -decoder” intends to recover messages for users in \mathcal{D} while regarding signals from users not in \mathcal{D} as interference, if the communication rate vector is within the operation region $\mathcal{R}_{\mathcal{D}}$. We define the following error probabilities for a $(\mathcal{D}, \mathcal{R}_{\mathcal{D}})$ -decoder. Conditioned on users in \mathcal{D} transmitting $(\mathbf{w}_{\mathcal{D}}, \mathbf{r}_{\mathcal{D}})$ and users not in \mathcal{D} choosing rate $\mathbf{r}_{\bar{\mathcal{D}}}$, let us denote the estimated messages and rates by $(\hat{\mathbf{w}}_{\mathcal{D}}, \hat{\mathbf{r}}_{\mathcal{D}})$ and $\hat{\mathbf{r}}_{\bar{\mathcal{D}}}$ if the decoder does not report a collision. We define the decoding error probability of the $(\mathcal{D}, \mathcal{R}_{\mathcal{D}})$ -decoder for $(\mathbf{w}_{\mathcal{D}}, \mathbf{r}_{\mathcal{D}}, \mathbf{r}_{\bar{\mathcal{D}}})$ with $\mathbf{r} \in \mathcal{R}_{\mathcal{D}}$ as

$$P_e(\mathbf{w}_{\mathcal{D}}, \mathbf{r}_{\mathcal{D}}, \mathbf{r}_{\bar{\mathcal{D}}}) = Pr\{(\hat{\mathbf{w}}_{\mathcal{D}}, \hat{\mathbf{r}}_{\mathcal{D}}) \neq (\mathbf{w}_{\mathcal{D}}, \mathbf{r}_{\mathcal{D}}) | (\mathbf{w}_{\mathcal{D}}, \mathbf{r}_{\mathcal{D}}, \mathbf{r}_{\bar{\mathcal{D}}})\}, \quad \forall (\mathbf{w}_{\mathcal{D}}, \mathbf{r}_{\mathcal{D}}, \mathbf{r}_{\bar{\mathcal{D}}}), \mathbf{r} \in \mathcal{R}_{\mathcal{D}}. \quad (11)$$

We define the collision miss detection probability for $(\mathbf{w}_{\mathcal{D}}, \mathbf{r}_{\mathcal{D}}, \mathbf{r}_{\bar{\mathcal{D}}})$ with $\mathbf{r} \notin \mathcal{R}_{\mathcal{D}}$ as

$$\bar{P}_c(\mathbf{w}_{\mathcal{D}}, \mathbf{r}_{\mathcal{D}}, \mathbf{r}_{\bar{\mathcal{D}}}) = 1 - Pr\{\text{“collision”} | (\mathbf{w}_{\mathcal{D}}, \mathbf{r}_{\mathcal{D}}, \mathbf{r}_{\bar{\mathcal{D}}})\} - Pr\{(\hat{\mathbf{w}}_{\mathcal{D}}, \hat{\mathbf{r}}_{\mathcal{D}}) = (\mathbf{w}_{\mathcal{D}}, \mathbf{r}_{\mathcal{D}}) | (\mathbf{w}_{\mathcal{D}}, \mathbf{r}_{\mathcal{D}}, \mathbf{r}_{\bar{\mathcal{D}}})\}, \\ \forall (\mathbf{w}_{\mathcal{D}}, \mathbf{r}_{\mathcal{D}}, \mathbf{r}_{\bar{\mathcal{D}}}), \mathbf{r} \notin \mathcal{R}_{\mathcal{D}}. \quad (12)$$

System error probability of the $(\mathcal{D}, \mathcal{R}_{\mathcal{D}})$ -decoder is defined by

$$P_{es}(\mathcal{D}, \mathcal{R}_{\mathcal{D}}) = \max \left\{ \max_{(\mathbf{w}_{\mathcal{D}}, \mathbf{r}_{\mathcal{D}}, \mathbf{r}_{\bar{\mathcal{D}}}), \mathbf{r} \in \mathcal{R}_{\mathcal{D}}} P_e(\mathbf{w}_{\mathcal{D}}, \mathbf{r}_{\mathcal{D}}, \mathbf{r}_{\bar{\mathcal{D}}}), \max_{(\mathbf{w}_{\mathcal{D}}, \mathbf{r}_{\mathcal{D}}, \mathbf{r}_{\bar{\mathcal{D}}}), \mathbf{r} \notin \mathcal{R}_{\mathcal{D}}} \bar{P}_c(\mathbf{w}_{\mathcal{D}}, \mathbf{r}_{\mathcal{D}}, \mathbf{r}_{\bar{\mathcal{D}}}) \right\}. \quad (13)$$

Given a finite codeword length N , the following lemma gives an upper bound on the achievable system error probability of a $(\mathcal{D}, \mathcal{R}_{\mathcal{D}})$ -decoder.

Lemma 1: The following system error probability bound is achievable for a K -user random multiple

access communication system over a discrete-time memoryless channel $P_{Y|X}$ with an $(\mathcal{D}, \mathcal{R}_\mathcal{D})$ -decoder,

$$P_{es}(\mathcal{D}, \mathcal{R}_\mathcal{D}) \leq \max \left\{ \max_{\mathbf{r} \in \mathcal{R}_\mathcal{D}} \sum_{\mathcal{S} \subset \mathcal{D}} \left[\sum_{\substack{\tilde{\mathbf{r}} \in \mathcal{R}_\mathcal{D}, \\ \tilde{\mathbf{r}}_\mathcal{S} = \mathbf{r}_\mathcal{S}}} \exp\{-NE_{m\mathcal{D}}(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}})\} + \max_{\substack{\mathbf{r}' \notin \mathcal{R}_\mathcal{D}, \\ \mathbf{r}'_\mathcal{S} = \mathbf{r}_\mathcal{S}}} \exp\{-NE_{i\mathcal{D}}(\mathcal{S}, \mathbf{r}, \mathbf{r}')\} \right], \right. \\ \left. \max_{\tilde{\mathbf{r}} \notin \mathcal{R}_\mathcal{D}} \sum_{\mathcal{S} \subset \mathcal{D}} \sum_{\substack{\mathbf{r} \in \mathcal{R}_\mathcal{D}, \\ \mathbf{r}_\mathcal{S} = \tilde{\mathbf{r}}_\mathcal{S}}} \max_{\substack{\mathbf{r}' \notin \mathcal{R}_\mathcal{D}, \\ \mathbf{r}'_\mathcal{S} = \tilde{\mathbf{r}}_\mathcal{S}}} \exp\{-NE_{i\mathcal{D}}(\mathcal{S}, \mathbf{r}, \mathbf{r}')\} \right\}, \quad (14)$$

where $E_{m\mathcal{D}}(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}})$ and $E_{i\mathcal{D}}(\mathcal{S}, \mathbf{r}, \mathbf{r}')$ are given by,

$$E_{m\mathcal{D}}(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}}) = \max_{0 < \rho \leq 1} -\rho \sum_{k \in \mathcal{D} \setminus \mathcal{S}} \tilde{r}_k + \max_{0 < s \leq 1} -\log \sum_Y \sum_{\mathbf{X}_\mathcal{S}} \prod_{k \in \mathcal{S}} P_{X|r_k}(X_k) \\ \times \left(\sum_{\mathbf{X}_{\mathcal{D} \setminus \mathcal{S}}} \prod_{k \in \mathcal{D} \setminus \mathcal{S}} P_{X|r_k}(X_k) P(Y|\mathbf{X}_\mathcal{D}, \mathbf{r}_{\bar{\mathcal{D}}})^{1-s} \right) \left(\sum_{\mathbf{X}_{\mathcal{D} \setminus \mathcal{S}}} \prod_{k \in \mathcal{D} \setminus \mathcal{S}} P_{X|\tilde{r}_k}(X_k) P(Y|\mathbf{X}_\mathcal{D}, \tilde{\mathbf{r}}_{\bar{\mathcal{D}}})^{\frac{s}{\rho}} \right)^\rho, \\ E_{i\mathcal{D}}(\mathcal{S}, \mathbf{r}, \mathbf{r}') = \max_{0 < \rho \leq 1} -\rho \sum_{k \in \mathcal{D} \setminus \mathcal{S}} r_k + \max_{0 < s \leq 1-\rho} -\log \sum_Y \sum_{\mathbf{X}_\mathcal{S}} \prod_{k \in \mathcal{S}} P_{X|r_k}(X_k) \\ \times \left(\sum_{\mathbf{X}_{\mathcal{D} \setminus \mathcal{S}}} \prod_{k \in \mathcal{D} \setminus \mathcal{S}} P_{X|r_k}(X_k) P(Y|\mathbf{X}_\mathcal{D}, \mathbf{r}_{\bar{\mathcal{D}}})^{\frac{s}{s+\rho}} \right)^{s+\rho} \left(\sum_{\mathbf{X}_{\mathcal{D} \setminus \mathcal{S}}} \prod_{k \in \mathcal{D} \setminus \mathcal{S}} P_{X|r'_k}(X_k) P(Y|\mathbf{X}_\mathcal{D}, \mathbf{r}'_{\bar{\mathcal{D}}}) \right)^{1-s}, \quad (15)$$

with $P(Y|\mathbf{X}_\mathcal{D}, \mathbf{r}_{\bar{\mathcal{D}}})$ in the above equations defined as

$$P(Y|\mathbf{X}_\mathcal{D}, \mathbf{r}_{\bar{\mathcal{D}}}) = \sum_{\mathbf{X}_{\bar{\mathcal{D}}}} \prod_{k \in \bar{\mathcal{D}}} P_{X|r_k}(X_k) P_{Y|\mathbf{X}}(Y|\mathbf{X}). \quad (16)$$

■

Proof: Since the decoder regards signals from users not in \mathcal{D} as interference, given that users not in \mathcal{D} choose rate $\mathbf{r}_{\bar{\mathcal{D}}}$, the multiple access channel experienced by users in \mathcal{D} is characterized by $P(Y|\mathbf{X}_\mathcal{D}, \mathbf{r}_{\bar{\mathcal{D}}})$ as specified in (16). The system can therefore be regarded as a random multiple access system with $|\mathcal{D}|$ users communicating over a compound channel characterized by the set $\{P(Y|\mathbf{X}_\mathcal{D}, \mathbf{r}_{\bar{\mathcal{D}}})|\forall \mathbf{r}_{\bar{\mathcal{D}}}\}$. Consequently, Lemma 1 is implied directly by Theorem 1. ■

Next, we will come back to the system where the receiver is only interested in the message of user k . We assume that for each user subset $\mathcal{D} \subseteq \{1, \dots, K\}$ with $k \in \mathcal{D}$, the receiver assigns an operation region $\mathcal{R}_\mathcal{D} \subseteq \mathcal{R}$ for the $(\mathcal{D}, \mathcal{R}_\mathcal{D})$ -decoder. That is, if the transmission rate \mathbf{r} satisfies $\mathbf{r} \in \mathcal{R}_\mathcal{D}$, the receiver intends to use the $(\mathcal{D}, \mathcal{R}_\mathcal{D})$ -decoder to recover the message of user k . It is easy to see that we should have,

$$\mathcal{R} = \bigcup_{\mathcal{D}: \mathcal{D} \subseteq \{1, \dots, K\}, k \in \mathcal{D}} \mathcal{R}_\mathcal{D}. \quad (17)$$

Assume that the receiver (single-user decoder) carries out all the $(\mathcal{D}, \mathcal{R}_\mathcal{D})$ -decoding operations. The receiver outputs an estimated message \hat{w}_k for user k if at least one $(\mathcal{D}, \mathcal{R}_\mathcal{D})$ -decoder outputs an estimated

message, and all estimation outputs of the $(\mathcal{D}, \mathcal{R}_{\mathcal{D}})$ -decoders for user k are identical. Otherwise, the receiver reports a collision for user k .

Let the transmitted rate vector be \mathbf{r} , and the transmitted message of user k be w_k . We define the decoding error probability $P_e(w_k, \mathbf{r})$, the collision miss detection probability $\bar{P}_c(w_k, \mathbf{r})$ and the system error probability P_{es} as follows,

$$\begin{aligned} P_e(w_k, \mathbf{r}) &= \Pr \{(\hat{w}_k, \hat{r}_k) \neq (w_k, r_k) | (w_k, \mathbf{r})\}, \forall (w_k, \mathbf{r}), \mathbf{r} \in \mathcal{R}, \\ \bar{P}_c(w_k, \mathbf{r}) &= 1 - \Pr \{ \text{"collision"} | (w_k, \mathbf{r})\} - \Pr \{(\hat{w}_k, \hat{r}_k) = (w_k, r_k) | (w_k, \mathbf{r})\}, \forall (w_k, \mathbf{r}), \mathbf{r} \notin \mathcal{R}, \\ P_{es} &= \max \left\{ \max_{(w_k, \mathbf{r}), \mathbf{r} \in \mathcal{R}} P_e(w_k, \mathbf{r}), \max_{(w_k, \mathbf{r}), \mathbf{r} \notin \mathcal{R}} \bar{P}_c(w_k, \mathbf{r}) \right\}. \end{aligned} \quad (18)$$

The following theorem gives an upper bound on the achievable system error probability of the single-user decoder.

Theorem 3: Consider a K -user random multiple access system over a discrete-time memoryless channel $P_{Y|X}$, with the receiver only interested in recovering the message for user k . Assume the receiver chooses an operation region \mathcal{R} . Let σ be an arbitrary partitioning of the operation region \mathcal{R} satisfying

$$\begin{aligned} \mathcal{R} &= \bigcup_{\mathcal{D}: \mathcal{D} \subseteq \{1, \dots, K\}, k \in \mathcal{D}} \mathcal{R}_{\mathcal{D}}, \\ \mathcal{R}_{\mathcal{D}'} \cap \mathcal{R}_{\mathcal{D}} &= \emptyset, \forall \mathcal{D}, \mathcal{D}' \subseteq \{1, \dots, K\}, \mathcal{D}' \neq \mathcal{D}, k \in \mathcal{D}, \mathcal{D}'. \end{aligned} \quad (19)$$

System error probability of the single-user decoder is upper-bounded by,

$$P_{es} \leq \min_{\sigma} \sum_{\mathcal{D}: \mathcal{D} \subseteq \{1, \dots, K\}, k \in \mathcal{D}} P_{es}(\mathcal{D}, \mathcal{R}_{\mathcal{D}}), \quad (20)$$

where $P_{es}(\mathcal{D}, \mathcal{R}_{\mathcal{D}})$ is the system error probability bound of the $(\mathcal{D}, \mathcal{R}_{\mathcal{D}})$ -decoder, and can be further bounded by (14). ■

Proof: Because a $(\mathcal{D}, \mathcal{R}_{\mathcal{D}})$ -decoder can always choose to report a collision even if it can decode the messages, its system error probability can be improved by shrinking the operation region $\mathcal{R}_{\mathcal{D}}$. This implies that the receiver of the random access system should partition its operation region \mathcal{R} into $\mathcal{R}_{\mathcal{D}}$ regions that do not overlap with each other. In other words, replacing (17) by (19) will improve the system error performance. The rest of the proof is implied by Lemma 1. ■

Note that the system error probability bound provided in Theorem 3 is implicit since the optimal partitioning scheme σ that maximize the right hand side of (20) is not specified. To find the optimal partitioning, one essentially needs to compute every single term on the right hand side of (20) and (14) for all rate options and all user subsets. Because both $E_{m\mathcal{D}}(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}})$ and $E_{i\mathcal{D}}(\mathcal{S}, \mathbf{r}, \mathbf{r}')$ defined in (15) involve the combinations of two user subsets and two rate vectors, the computational complexity of finding the optimal partitioning scheme is $O((2M)^{2K})$.

IV. CONCLUSIONS

We investigated the error performance of the random multiple access system over a compound discrete-time memoryless channel. An achievable bound on the system error probability was derived under the non-asymptotic assumption of a finite codeword length. We showed that the results can be extended to the random multiple access system where the receiver is only interested in decoding messages for a user subset.

APPENDIX

A. Proof of Theorem 1

We assume that the following decoding algorithm is used at the receiver. Given the received channel output symbols \mathbf{y} , the receiver outputs a message and rate vector pair (\mathbf{w}, \mathbf{r}) together with a channel realization $P_{Y|X}$ such that $(\mathbf{r}, P_{Y|X}) \in \mathcal{R}$ if the following condition is satisfied, for all user subsets $\mathcal{S} \subset \{1, \dots, K\}$,

$$\begin{aligned} & -\frac{1}{N} \log \Pr\{\mathbf{y}|\mathbf{x}_{(\mathbf{w}, \mathbf{r})}, P_{Y|X}\} < -\frac{1}{N} \log \Pr\{\mathbf{y}|\mathbf{x}_{(\tilde{\mathbf{w}}, \tilde{\mathbf{r}})}, \tilde{P}_{Y|X}\}, \\ & \text{for all } (\tilde{\mathbf{w}}, \tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\tilde{\mathbf{w}}_{\mathcal{S}}, \tilde{\mathbf{r}}_{\mathcal{S}}) = (\mathbf{w}_{\mathcal{S}}, \mathbf{r}_{\mathcal{S}}), (\tilde{w}_k, \tilde{r}_k) \neq (w_k, r_k), \forall k \notin \mathcal{S}, \\ & \text{and } (\tilde{\mathbf{w}}, \tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\mathbf{w}, \mathbf{r}, P_{Y|X}) \in \mathcal{R}_{(\mathcal{S}, \mathbf{y})}, \text{ with} \\ & \mathcal{R}_{(\mathcal{S}, \mathbf{y})} = \left\{ (\tilde{\mathbf{w}}, \tilde{\mathbf{r}}, \tilde{P}_{Y|X}) \mid (\tilde{\mathbf{r}}, \tilde{P}_{Y|X}) \in \mathcal{R}, -\frac{1}{N} \log \Pr\{\mathbf{y}|\mathbf{x}_{(\tilde{\mathbf{w}}, \tilde{\mathbf{r}})}, \tilde{P}_{Y|X}\} < \tau_{(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}, \mathcal{S})}(\mathbf{y}) \right\}, \quad (21) \end{aligned}$$

where $\tau_{(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}, \mathcal{S})}(\cdot)$ is a per-determined typicality threshold function of the channel output symbols \mathbf{y} , associated with the rate and channel realization pair $(\tilde{\mathbf{r}}, \tilde{P}_{Y|X})$ and the user subset \mathcal{S} . If there is no codeword satisfying (21), the receiver reports a collision. In other words, for a given \mathcal{S} , the receiver searches for the subset of codewords with likelihood values larger than the corresponding typicality threshold. If the subset is not empty, the receiver outputs the codeword with the maximum likelihood value as the estimate for this given \mathcal{S} . If the estimates for all $\mathcal{S} \subset \{1, \dots, K\}$ agree with each other, the receiver regards this estimate as the decoding decision and outputs the corresponding decoded message and rate pair. Otherwise, the receiver reports a collision. Note that in (21), for given \mathcal{S} and (\mathbf{w}, \mathbf{r}) , we only compare the likelihood value of codeword vector $\mathbf{x}_{(\mathbf{w}, \mathbf{r})}$ with those of the codeword vectors satisfying $(\tilde{\mathbf{w}}_{\mathcal{S}}, \tilde{\mathbf{r}}_{\mathcal{S}}) = (\mathbf{w}_{\mathcal{S}}, \mathbf{r}_{\mathcal{S}}), (\tilde{w}_k, \tilde{r}_k) \neq (w_k, r_k), \forall k \notin \mathcal{S}$. We will first analyze the error performance for each user subset \mathcal{S} and then derive the overall error performance by taking the union over all \mathcal{S} .

Given a user subset $\mathcal{S} \subset \{1, \dots, K\}$, we define the following probability terms.

First, assume (\mathbf{w}, \mathbf{r}) is transmitted over channel $P_{Y|X}$, with $(\mathbf{r}, P_{Y|X}) \in \mathcal{R}$. Let $P_{t[\mathbf{r}, P_{Y|X}, \mathcal{S}]}$ be the probability that the likelihood value of the transmitted codeword vector over the channel $P_{Y|X}$ is no

larger than the corresponding typicality threshold,

$$P_{t[(\mathbf{r}, P_{Y|X}), S]} = Pr \left\{ P(\mathbf{y}|\mathbf{x}_{(w, r)}, P_{Y|X}) \leq e^{-N\tau_{(\mathbf{r}, P_{Y|X}, S)}(\mathbf{y})} \right\}. \quad (22)$$

Define $P_{m[(\mathbf{r}, P_{Y|X}), (\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), S]}$ as the probability that the likelihood value of the transmitted codeword vector over the channel realization $P_{Y|X}$ is no larger than that of another codeword $(\tilde{\mathbf{w}}, \tilde{\mathbf{r}})$ with $(\tilde{\mathbf{w}}_S, \tilde{\mathbf{r}}_S) = (\mathbf{w}_S, \mathbf{r}_S)$, $(\tilde{w}_k, \tilde{r}_k) \neq (w_k, r_k), \forall k \notin S$, over channel $\tilde{P}_{Y|X}$ with $(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}) \in \mathcal{R}$,

$$P_{m[(\mathbf{r}, P_{Y|X}), (\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), S]} = Pr \left\{ P(\mathbf{y}|\mathbf{x}_{(w, r)}, P_{Y|X}) \leq P(\mathbf{y}|\mathbf{x}_{(\tilde{\mathbf{w}}, \tilde{\mathbf{r}})}, \tilde{P}_{Y|X}) \right\} \\ (\tilde{\mathbf{w}}, \tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\tilde{\mathbf{r}}, \tilde{P}_{Y|X}) \in \mathcal{R}, (\tilde{\mathbf{w}}_S, \tilde{\mathbf{r}}_S) = (\mathbf{w}_S, \mathbf{r}_S), (\tilde{w}_k, \tilde{r}_k) \neq (w_k, r_k), \forall k \notin S. \quad (23)$$

Second, assume that $(\tilde{\mathbf{w}}, \tilde{\mathbf{r}})$ is transmitted over channel $\tilde{P}_{Y|X}$, with $(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}) \notin \mathcal{R}$. Define $P_{i[(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\mathbf{r}, P_{Y|X}), S]}$ as the probability that the decoder finds a codeword (\mathbf{w}, \mathbf{r}) with $(\mathbf{w}_S, \mathbf{r}_S) = (\tilde{\mathbf{w}}_S, \tilde{\mathbf{r}}_S)$, $(w_k, r_k) \neq (\tilde{w}_k, \tilde{r}_k), \forall k \notin S$, over channel $P_{Y|X}$ with $(\mathbf{r}, P_{Y|X}) \in \mathcal{R}$, such that its likelihood value is larger than the corresponding typicality threshold,

$$P_{i[(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\mathbf{r}, P_{Y|X}), S]} = Pr \left\{ P(\mathbf{y}|\mathbf{x}_{(w, r)}, P_{Y|X}) > e^{-N\tau_{(\mathbf{r}, P_{Y|X}, S)}(\mathbf{y})} \right\}, \\ (\mathbf{w}, \mathbf{r}, P_{Y|X}), (\mathbf{r}, P_{Y|X}) \in \mathcal{R}, (\mathbf{w}_S, \mathbf{r}_S) = (\tilde{\mathbf{w}}_S, \tilde{\mathbf{r}}_S), (w_k, r_k) \neq (\tilde{w}_k, \tilde{r}_k), \forall k \notin S. \quad (24)$$

With the above probability definitions, by applying the union bound over all S , we can upper-bound the system error probability by

$$P_{es} \leq \max \left\{ \begin{array}{l} \max_{(\mathbf{r}, P_{Y|X}) \in \mathcal{R}} \sum_{S \subset \{1, \dots, K\}} \left[P_{t[(\mathbf{r}, P_{Y|X}), S]} + \sum_{(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}) \in \mathcal{R}, \tilde{\mathbf{r}}_S = \mathbf{r}_S} P_{m[(\mathbf{r}, P_{Y|X}), (\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), S]} \right], \\ \max_{(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}) \notin \mathcal{R}} \sum_{S \subset \{1, \dots, K\}} \sum_{(\mathbf{r}, P_{Y|X}) \in \mathcal{R}, \mathbf{r}_S = \tilde{\mathbf{r}}_S} P_{i[(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\mathbf{r}, P_{Y|X}), S]} \end{array} \right\}. \quad (25)$$

Next, we will derive individual upper-bounds for each of the probability terms on the right hand side of (25).

Step I: Upper-bounding $P_{m[(\mathbf{r}, P_{Y|X}), (\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), S]}$

Denote E_θ as the expectation operator over random variable θ which is defined in Section II. Consequently, given $(\mathbf{r}, P_{Y|X}), (\tilde{\mathbf{r}}, \tilde{P}_{Y|X}) \in \mathcal{R}$, $P_{m[(\mathbf{r}, P_{Y|X}), (\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), S]}$ defined in (23) can be rewritten as

$$P_{m[(\mathbf{r}, P_{Y|X}), (\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), S]} = E_\theta \left[\sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}_{(w, r)}, P_{Y|X}) \phi_{m[(\mathbf{r}, P_{Y|X}), (\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), S]}(\mathbf{y}) \right], \quad (26)$$

where $\phi_{m[(\mathbf{r}, P_{Y|X}), (\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), S]}(\mathbf{y}) = 1$ if $P(\mathbf{y}|\mathbf{x}_{(w, r)}, P_{Y|X}) \leq P(\mathbf{y}|\mathbf{x}_{(\tilde{\mathbf{w}}, \tilde{\mathbf{r}})}, \tilde{P}_{Y|X})$ for some triplet $(\tilde{\mathbf{w}}, \tilde{\mathbf{r}}, \tilde{P}_{Y|X})$ with $(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}) \in \mathcal{R}$, $(\tilde{\mathbf{w}}_S, \tilde{\mathbf{r}}_S) = (\mathbf{w}_S, \mathbf{r}_S)$, $(\tilde{w}_k, \tilde{r}_k) \neq (w_k, r_k), \forall k \notin S$. Otherwise, $\phi_{m[(\mathbf{r}, P_{Y|X}), (\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), S]}(\mathbf{y}) = 0$. We can upper-bound $\phi_{m[(\mathbf{r}, P_{Y|X}), (\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), S]}(\mathbf{y})$ for any constants $\rho > 0$ and $s > 0$ as follows,

$$\phi_{m[(\mathbf{r}, P_{Y|X}), (\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), S]}(\mathbf{y}) \leq \left[\frac{\sum_{\tilde{\mathbf{w}}, (\tilde{\mathbf{w}}_S, \tilde{\mathbf{r}}_S) = (\mathbf{w}_S, \mathbf{r}_S), (\tilde{w}_k, \tilde{r}_k) \neq (w_k, r_k), \forall k \notin S} P(\mathbf{y}|\mathbf{x}_{(\tilde{\mathbf{w}}, \tilde{\mathbf{r}})}, \tilde{P}_{Y|X})^{\frac{s}{\rho}}}{P(\mathbf{y}|\mathbf{x}_{(w, r)}, P_{Y|X})^{\frac{s}{\rho}}} \right]^\rho. \quad (27)$$

Substituting (27) back into (26) gives,

$$\begin{aligned}
P_{m[(\mathbf{r}, P_{Y|\mathbf{X}}), (\tilde{\mathbf{r}}, \tilde{P}_{Y|\mathbf{X}}), \mathcal{S}]} &\leq E_\theta \left[\sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}(\mathbf{w}, \mathbf{r}), P_{Y|\mathbf{X}}) \right. \\
&\quad \times \left. \left[\frac{\sum_{\tilde{\mathbf{w}}, (\tilde{\mathbf{w}}_S, \tilde{\mathbf{r}}_S) = (\mathbf{w}_S, \mathbf{r}_S), (\tilde{\mathbf{w}}_k, \tilde{\mathbf{r}}_k) \neq (\mathbf{w}_k, \mathbf{r}_k), \forall k \notin \mathcal{S}} P(\mathbf{y}|\mathbf{x}(\tilde{\mathbf{w}}, \tilde{\mathbf{r}}), \tilde{P}_{Y|\mathbf{X}})^{\frac{s}{\rho}}}{P(\mathbf{y}|\mathbf{x}(\mathbf{w}, \mathbf{r}), P_{Y|\mathbf{X}})^{\frac{s}{\rho}}} \right]^\rho \right] \\
&= \sum_{\mathbf{y}} E_{\theta_S} \left[E_{\theta_{\bar{S}}} \left[P(\mathbf{y}|\mathbf{x}(\mathbf{w}, \mathbf{r}), P_{Y|\mathbf{X}})^{1-s} \right] \right. \\
&\quad \times E_{\theta_{\bar{S}}} \left[\left[\sum_{\tilde{\mathbf{w}}, (\tilde{\mathbf{w}}_S, \tilde{\mathbf{r}}_S) = (\mathbf{w}_S, \mathbf{r}_S), (\tilde{\mathbf{w}}_k, \tilde{\mathbf{r}}_k) \neq (\mathbf{w}_k, \mathbf{r}_k), \forall k \notin \mathcal{S}} P(\mathbf{y}|\mathbf{x}(\tilde{\mathbf{w}}, \tilde{\mathbf{r}}), \tilde{P}_{Y|\mathbf{X}})^{\frac{s}{\rho}} \right]^\rho \right] \left. \right]. \quad (28)
\end{aligned}$$

The second step in (28) is due to the independence between the codewords corresponding to $(\mathbf{w}_{\bar{S}}, \mathbf{r}_{\bar{S}})$ and $(\tilde{\mathbf{w}}_{\bar{S}}, \tilde{\mathbf{r}}_{\bar{S}})$.

With the assumption of $0 < \rho \leq 1$, we can further bound $P_{m[(\mathbf{r}, P_{Y|\mathbf{X}}), (\tilde{\mathbf{r}}, \tilde{P}_{Y|\mathbf{X}}), \mathcal{S}]}$ by

$$\begin{aligned}
P_{m[(\mathbf{r}, P_{Y|\mathbf{X}}), (\tilde{\mathbf{r}}, \tilde{P}_{Y|\mathbf{X}}), \mathcal{S}]} &\leq \sum_{\mathbf{y}} E_{\theta_S} \left[E_{\theta_{\bar{S}}} \left[P(\mathbf{y}|\mathbf{x}(\mathbf{w}, \mathbf{r}), P_{Y|\mathbf{X}})^{1-s} \right] \right. \\
&\quad \times E_{\theta_{\bar{S}}} \left[\left[\sum_{\tilde{\mathbf{w}}, (\tilde{\mathbf{w}}_S, \tilde{\mathbf{r}}_S) = (\mathbf{w}_S, \mathbf{r}_S)} P(\mathbf{y}|\mathbf{x}(\tilde{\mathbf{w}}, \tilde{\mathbf{r}}), \tilde{P}_{Y|\mathbf{X}})^{\frac{s}{\rho}} \right]^\rho \right] \left. \right] \\
&\leq e^{N\rho \sum_{k \notin \mathcal{S}} \tilde{r}_k} \sum_{\mathbf{y}} E_{\theta_S} \left[E_{\theta_{\bar{S}}} \left[P(\mathbf{y}|\mathbf{x}(\mathbf{w}, \mathbf{r}), P_{Y|\mathbf{X}})^{1-s} \right] E_{\theta_{\bar{S}}} \left[\left[P(\mathbf{y}|\mathbf{x}(\tilde{\mathbf{w}}, \tilde{\mathbf{r}}), \tilde{P}_{Y|\mathbf{X}})^{\frac{s}{\rho}} \right]^\rho \right] \right]. \quad (29)
\end{aligned}$$

It is easy to verify that the bound in (29) holds for all $0 < \rho \leq 1$ and $s > 0$, and becomes trivial for $s > 1$. Consequently, (29) gives the following upper bound,

$$P_{m[(\mathbf{r}, P_{Y|\mathbf{X}}), (\tilde{\mathbf{r}}, \tilde{P}_{Y|\mathbf{X}}), \mathcal{S}]} \leq \exp \left\{ -N E_m(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}}, P_{Y|\mathbf{X}}, \tilde{P}_{Y|\mathbf{X}}) \right\}, \quad (30)$$

where $E_m(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}}, P_{Y|\mathbf{X}}, \tilde{P}_{Y|\mathbf{X}})$ is specified in (5).

Step II: Upper-bounding $P_{t[\mathbf{r}, P_{Y|\mathbf{X}}, \mathcal{S}]}$

Given that $(\mathbf{r}, P_{Y|\mathbf{X}}) \in \mathcal{R}$, we can rewrite $P_{t[\mathbf{r}, P_{Y|\mathbf{X}}, \mathcal{S}]}$, defined in (22), as follows,

$$P_{t[\mathbf{r}, P_{Y|\mathbf{X}}, \mathcal{S}]} = E_\theta \left[\sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}(\mathbf{w}, \mathbf{r}), P_{Y|\mathbf{X}}) \phi_{t[\mathbf{r}, P_{Y|\mathbf{X}}, \mathcal{S}]}(\mathbf{y}) \right], \quad (31)$$

where $\phi_{t[\mathbf{r}, P_{Y|\mathbf{X}}, \mathcal{S}]}(\mathbf{y}) = 1$ if $P(\mathbf{y}|\mathbf{x}(\mathbf{w}, \mathbf{r}), P_{Y|\mathbf{X}}) \leq e^{-N\tau(\mathbf{r}, P_{Y|\mathbf{X}}, \mathcal{S})(\mathbf{y})}$, otherwise $\phi_{t[\mathbf{r}, P_{Y|\mathbf{X}}, \mathcal{S}]}(\mathbf{y}) = 0$. Note that the value of $\tau(\mathbf{r}, P_{Y|\mathbf{X}}, \mathcal{S})(\mathbf{y})$ will be determined in Step IV. Similarly, we can bound $\phi_{t[\mathbf{r}, P_{Y|\mathbf{X}}, \mathcal{S}]}(\mathbf{y})$, for any $s_1 > 0$, as follows,

$$\phi_{t[\mathbf{r}, P_{Y|\mathbf{X}}, \mathcal{S}]}(\mathbf{y}) \leq \frac{e^{-Ns_1\tau(\mathbf{r}, P_{Y|\mathbf{X}}, \mathcal{S})(\mathbf{y})}}{P(\mathbf{y}|\mathbf{x}(\mathbf{w}, \mathbf{r}), P_{Y|\mathbf{X}})^{s_1}}. \quad (32)$$

This yields,

$$\begin{aligned}
P_{t[\mathbf{r}, P_{Y|\mathbf{X}}, \mathcal{S}]} &\leq E_\theta \left[\sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}(\mathbf{w}, \mathbf{r}), P_{Y|\mathbf{X}})^{1-s_1} e^{-Ns_1\tau(\mathbf{r}, P_{Y|\mathbf{X}}, \mathcal{S})(\mathbf{y})} \right] \\
&= \sum_{\mathbf{y}} E_{\theta_S} \left[E_{\theta_{\bar{S}}} \left[P(\mathbf{y}|\mathbf{x}(\mathbf{w}, \mathbf{r}), P_{Y|\mathbf{X}})^{1-s_1} \right] e^{-Ns_1\tau(\mathbf{r}, P_{Y|\mathbf{X}}, \mathcal{S})(\mathbf{y})} \right]. \quad (33)
\end{aligned}$$

Step III: Upper-Bounding $P_{i[(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\mathbf{r}, P_{Y|X}), S]}$

Given $\tilde{\mathbf{r}} \notin \mathcal{R}$ and $\mathbf{r} \in \mathcal{R}$, we rewrite $P_{i[(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\mathbf{r}, P_{Y|X}), S]}$ as

$$P_{i[(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\mathbf{r}, P_{Y|X}), S]} = E_{\theta} \left[\sum_{\mathbf{y}} P(\mathbf{y}|\mathbf{x}(\tilde{\mathbf{w}}, \tilde{\mathbf{r}}), \tilde{P}_{Y|X}) \phi_{[(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\mathbf{r}, P_{Y|X}), S]}(\mathbf{y}) \right], \quad (34)$$

where $\phi_{[(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\mathbf{r}, P_{Y|X}), S]}(\mathbf{y}) = 1$ if there exists a triplet $(\mathbf{w}, \mathbf{r}, P_{Y|X})$ with $(\mathbf{r}, P_{Y|X}) \in \mathcal{R}$, $(\mathbf{w}_S, \mathbf{r}_S) = (\tilde{\mathbf{w}}_S, \tilde{\mathbf{r}}_S)$, and $(w_k, r_k) \neq (\tilde{w}_k, \tilde{r}_k)$ for all $k \notin S$, such that $P(\mathbf{y}|\mathbf{x}(\mathbf{w}, \mathbf{r}), P_{Y|X}) > e^{-N\tau(\mathbf{r}, P_{Y|X}, S)(\mathbf{y})}$ is satisfied. Otherwise, $\phi_{[(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\mathbf{r}, P_{Y|X}), S]}(\mathbf{y}) = 0$.

For any $s_2 > 0$ and $\tilde{\rho} > 0$, $\phi_{[(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\mathbf{r}, P_{Y|X}), S]}(\mathbf{y})$ can be bounded by,

$$\phi_{[(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\mathbf{r}, P_{Y|X}), S]}(\mathbf{y}) \leq \left[\frac{\sum_{\mathbf{w}, (\mathbf{w}_S, \mathbf{r}_S) = (\tilde{\mathbf{w}}_S, \tilde{\mathbf{r}}_S), (w_k, r_k) \neq (\tilde{w}_k, \tilde{r}_k) \forall k \notin S} P(\mathbf{y}|\mathbf{x}(\mathbf{w}, \mathbf{r}), P_{Y|X})^{\frac{s_2}{\tilde{\rho}}}}{e^{-N\frac{s_2}{\tilde{\rho}}\tau(\mathbf{r}, P_{Y|X}, S)(\mathbf{y})}} \right]^{\tilde{\rho}}. \quad (35)$$

Substituting (35) into (34) yields,

$$\begin{aligned} P_{i[(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\mathbf{r}, P_{Y|X}), S]} &\leq \sum_{\mathbf{y}} E_{\theta} \left[P(\mathbf{y}|\mathbf{x}(\tilde{\mathbf{w}}, \tilde{\mathbf{r}}), \tilde{P}_{Y|X}) e^{Ns_2\tau(\mathbf{r}, P_{Y|X}, S)(\mathbf{y})} \right. \\ &\quad \times \left. \left[\sum_{\mathbf{w}, (\mathbf{w}_S, \mathbf{r}_S) = (\tilde{\mathbf{w}}_S, \tilde{\mathbf{r}}_S)} P(\mathbf{y}|\mathbf{x}(\mathbf{w}, \mathbf{r}), P_{Y|X})^{\frac{s_2}{\tilde{\rho}}} \right]^{\tilde{\rho}} \right]. \end{aligned} \quad (36)$$

The independence between $(\mathbf{w}_{\bar{S}}, \mathbf{r}_{\bar{S}})$ and $(\tilde{\mathbf{w}}_{\bar{S}}, \tilde{\mathbf{r}}_{\bar{S}})$ allows us to rewrite the above bound as

$$\begin{aligned} P_{i[(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\mathbf{r}, P_{Y|X}), S]} &\leq \sum_{\mathbf{y}} E_{\theta_S} \left[E_{\theta_{\bar{S}}} \left[P(\mathbf{y}|\mathbf{x}(\tilde{\mathbf{w}}, \tilde{\mathbf{r}}), \tilde{P}_{Y|X}) \right] e^{Ns_2\tau(\mathbf{r}, P_{Y|X}, S)(\mathbf{y})} \right. \\ &\quad \times E_{\theta_{\bar{S}}} \left[\left[\sum_{\mathbf{w}, (\mathbf{w}_S, \mathbf{r}_S) = (\tilde{\mathbf{w}}_S, \tilde{\mathbf{r}}_S)} P(\mathbf{y}|\mathbf{x}(\mathbf{w}, \mathbf{r}), P_{Y|X})^{\frac{s_2}{\tilde{\rho}}} \right]^{\tilde{\rho}} \right] \left. \right]. \end{aligned} \quad (37)$$

With the assumption of $0 < \tilde{\rho} \leq 1$, the inequality in (37) becomes

$$\begin{aligned} P_{i[(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\mathbf{r}, P_{Y|X}), S]} &\leq \sum_{\mathbf{y}} E_{\theta_S} \left[E_{\theta_{\bar{S}}} \left[P(\mathbf{y}|\mathbf{x}(\tilde{\mathbf{w}}, \tilde{\mathbf{r}}), \tilde{P}_{Y|X}) \right] \right. \\ &\quad \times E_{\theta_{\bar{S}}} \left\{ \left[P(\mathbf{y}|\mathbf{x}(\mathbf{w}, \mathbf{r}), P_{Y|X})^{\frac{s_2}{\tilde{\rho}}} \right]^{\tilde{\rho}} e^{Ns_2\tau(\mathbf{r}, P_{Y|X}, S)(\mathbf{y})} e^{N\tilde{\rho} \sum_{k \notin S} r_k} \right\} \\ &\leq \max_{(\mathbf{r}', P'_{Y|X}) \notin \mathcal{R}, \mathbf{r}'_S = \mathbf{r}_S} \sum_{\mathbf{y}} E_{\theta_S} \left[E_{\theta_{\bar{S}}} \left[P(\mathbf{y}|\mathbf{x}(\mathbf{w}', \mathbf{r}'), P'_{Y|X}) \right] \right. \\ &\quad \times E_{\theta_{\bar{S}}} \left\{ \left[P(\mathbf{y}|\mathbf{x}(\mathbf{w}, \mathbf{r}), P_{Y|X})^{\frac{s_2}{\tilde{\rho}}} \right]^{\tilde{\rho}} e^{Ns_2\tau(\mathbf{r}, P_{Y|X}, S)(\mathbf{y})} e^{N\tilde{\rho} \sum_{k \notin S} r_k} \right\} \left. \right]. \end{aligned} \quad (38)$$

Note that the upper bound in (38) is no longer a function of $(\tilde{\mathbf{r}}_{\bar{S}}, \tilde{P}_{Y|X})$.

Step IV: Choosing $\tau(\mathbf{r}, P_{Y|X}, S)(\mathbf{y})$

The value of $\tau(\mathbf{r}, P_{Y|X}, S)(\mathbf{y})$ can be determined by jointly optimizing the bounds in (33) and (38). Consequently, given $(\mathbf{r}, P_{Y|X}) \in \mathcal{R}$, \mathbf{y} and auxiliary variables $s_1 > 0$, $s_2 > 0$, $0 < \tilde{\rho} \leq 1$, we choose $\tau(\mathbf{r}, P_{Y|X}, S)(\mathbf{y})$ such that the following equality is satisfied,

$$\begin{aligned} &E_{\theta_{\bar{S}}} \left[P(\mathbf{y}|\mathbf{x}(\mathbf{w}, \mathbf{r}), P_{Y|X})^{1-s_1} \right] e^{-Ns_1\tau(\mathbf{r}, P_{Y|X}, S)(\mathbf{y})} \\ &= E_{\theta_{\bar{S}}} \left[P(\mathbf{y}|\mathbf{x}(\tilde{\mathbf{w}}^*, \tilde{\mathbf{r}}^*), \tilde{P}_{Y|X}) \right] \times E_{\theta_{\bar{S}}} \left\{ \left[P(\mathbf{y}|\mathbf{x}(\mathbf{w}, \mathbf{r}), P_{Y|X})^{\frac{s_2}{\tilde{\rho}}} \right]^{\tilde{\rho}} e^{Ns_2\tau(\mathbf{r}, P_{Y|X}, S)(\mathbf{y})} e^{N\tilde{\rho} \sum_{k \notin S} r_k} \right\}. \end{aligned} \quad (39)$$

where $(\tilde{\mathbf{r}}^*, \tilde{P}_{Y|X}^*)$ is defined as³

$$(\tilde{\mathbf{r}}^*, \tilde{P}_{Y|X}^*) = \underset{(\mathbf{r}', P'_{Y|X}) \notin \mathcal{R}, \mathbf{r}'_S = \mathbf{r}_S}{\operatorname{argmax}} \sum_{\mathbf{y}} E_{\theta_S} \left[E_{\theta_{\bar{S}}} \left[P(\mathbf{y} | \mathbf{x}_{(w', \mathbf{r}')}), P'_{Y|X} \right] \right. \\ \left. \times E_{\theta_{\bar{S}}} \left\{ \left[P(\mathbf{y} | \mathbf{x}_{(w, \mathbf{r})}), P_{Y|X} \right]^{\frac{s_2}{\tilde{\rho}}} \right\}^{\tilde{\rho}} e^{Ns_2 \tau_{(r, P_{Y|X}, S)}(\mathbf{y})} e^{N\tilde{\rho} \sum_{k \notin S} r_k} \right]. \quad (40)$$

Finding a solution for (39) is always possible since that the left hand side of (39) decreases with $\tau_{(r, P_{Y|X}, S)}(\mathbf{y})$, while the right hand side of (39) increases with $\tau_{(r, P_{Y|X}, S)}(\mathbf{y})$. This yields the desired typicality threshold, denoted by $\tau_{(r, P_{Y|X}, S)}^*(\mathbf{y})$, which gives

$$e^{-N\tau_{(r, P_{Y|X}, S)}^*(\mathbf{y})} = \frac{\left\{ E_{\theta_{\bar{S}}} \left[P(\mathbf{y} | \mathbf{x}_{(\tilde{w}^*, \tilde{\mathbf{r}}^*)}), \tilde{P}_{Y|X}^* \right] \right\}^{\frac{1}{s_1+s_2}} E_{\theta_{\bar{S}}} \left\{ \left[P(\mathbf{y} | \mathbf{x}_{(w, \mathbf{r})}), P_{Y|X} \right]^{\frac{s_2}{\tilde{\rho}}} \right\}^{\frac{\tilde{\rho}}{s_1+s_2}}}{\left\{ E_{\theta_{\bar{S}}} \left[P(\mathbf{y} | \mathbf{x}_{(w, \mathbf{r})}), P_{Y|X} \right]^{1-s_1} \right\}^{\frac{1}{s_1+s_2}}} \times e^{N\frac{\tilde{\rho}}{s_1+s_2} \sum_{k \notin S} r_k}. \quad (41)$$

Substituting (41) into (33), we get

$$P_{t[r, P_{Y|X}, S]} \leq \sum_{\mathbf{y}} E_{\theta_S} \left[E_{\theta_{\bar{S}}} \left[P(\mathbf{y} | \mathbf{x}_{(w, \mathbf{r})}), P_{Y|X} \right]^{1-s_1} \right]^{\frac{s_2}{s_1+s_2}} \left\{ E_{\theta_{\bar{S}}} \left[P(\mathbf{y} | \mathbf{x}_{(\tilde{w}^*, \tilde{\mathbf{r}}^*)}), \tilde{P}_{Y|X}^* \right] \right\}^{\frac{s_1}{s_1+s_2}} \\ \times E_{\theta_{\bar{S}}} \left\{ \left[P(\mathbf{y} | \mathbf{x}_{(w, \mathbf{r})}), P_{Y|X} \right]^{\frac{s_2}{\tilde{\rho}}} \right\}^{\frac{s_1 \tilde{\rho}}{s_1+s_2}} e^{N\frac{s_1 \tilde{\rho}}{s_1+s_2} \sum_{k \notin S} r_k} \right]. \quad (42)$$

Let $s_2 < \tilde{\rho}$ and $s_1 = 1 - \frac{s_2}{\tilde{\rho}}$, and then do a variable change with $\rho = \frac{\tilde{\rho}(\tilde{\rho}-s_2)}{\tilde{\rho}-(1-\tilde{\rho})s_2}$ and $s = 1 - \frac{\tilde{\rho}-s_2}{\tilde{\rho}-(1-\tilde{\rho})s_2}$. Consequently, inequality (42) becomes,

$$P_{t[r, P_{Y|X}, S]} \leq \max_{(\mathbf{r}', P'_{Y|X}) \notin \mathcal{R}, \mathbf{r}'_S = \mathbf{r}_S} e^{N\rho \sum_{k \notin S} r_k} \left\{ \sum_Y \sum_{\mathbf{X}_S} \prod_{k \in S} P_{X|r_k}(X_k) \right. \\ \left. \times \left(\sum_{\mathbf{X}_{\bar{S}}} \prod_{k \notin S} P_{X|r_k}(X_k) P_{Y|X}(Y|\mathbf{X})^{\frac{s}{s+\rho}} \right)^{s+\rho} \left(\sum_{\mathbf{X}_{\bar{S}}} \prod_{k \notin S} P_{X|r'_k}(X_k) P'_{Y|X}(Y|\mathbf{X}) \right)^{1-s} \right\}^N. \quad (43)$$

Similarly, we can obtain the same upper bound for $P_{i[(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\mathbf{r}, P_{Y|X}), S]}$ as given at the right hand side of (43). Since (43) holds for all $0 < \rho \leq 1$ and $0 < s \leq 1 - \rho$, we have

$$P_{t[r, P_{Y|X}, S]}, P_{i[(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}), (\mathbf{r}, P_{Y|X}), S]} \leq \max_{(\mathbf{r}', P'_{Y|X}) \notin \mathcal{R}, \mathbf{r}'_S = \mathbf{r}_S} \exp \left\{ -N E_i(\mathcal{S}, \mathbf{r}, \mathbf{r}', P_{Y|X}, P'_{Y|X}) \right\}, \quad (44)$$

where $E_i(\mathcal{S}, \mathbf{r}, \mathbf{r}', P_{Y|X}, P'_{Y|X})$ is given in (5).

By substituting (30) and (44) into (25), we get the desired result.

³Although the notation of $\tilde{\mathbf{w}}^*$ is used in (39), the result is actually invariant to any choice of the message vector.

B. Proof of Theorem 2

We assume that the following decoding algorithm is used at the receiver. Given the channel output sequence \mathbf{y} , the receiver outputs a message and rate vector pair (\mathbf{w}, \mathbf{r}) together with a channel class \mathcal{F} such that $(\mathbf{r}, \mathcal{F}) \in \mathcal{R}$ if for all user subset $\mathcal{S} \subset \{1, \dots, K\}$, the following condition is satisfied,

$$\begin{aligned} -\frac{1}{N} \log \Pr \left\{ \mathbf{y} | \mathbf{x}_{(\mathbf{w}, \mathbf{r})}, P_{\min}^{\mathcal{F}} \right\} &< -\frac{1}{N} \log \Pr \left\{ \mathbf{y} | \mathbf{x}_{(\tilde{\mathbf{w}}, \tilde{\mathbf{r}})}, P_{\max}^{\tilde{\mathcal{F}}} \right\}, \\ \text{for all } (\tilde{\mathbf{w}}, \tilde{\mathbf{r}}, \tilde{\mathcal{F}}), (\tilde{\mathbf{w}}_{\mathcal{S}}, \tilde{\mathbf{r}}_{\mathcal{S}}) &= (\mathbf{w}_{\mathcal{S}}, \mathbf{r}_{\mathcal{S}}), (\tilde{w}_k, \tilde{r}_k) \neq (w_k, r_k), \forall k \notin \mathcal{S}, \\ \text{and } (\mathbf{w}, \mathbf{r}, P_{\min}^{\mathcal{F}}), (\tilde{\mathbf{w}}, \tilde{\mathbf{r}}, P_{\max}^{\tilde{\mathcal{F}}}) &\in \mathcal{R}_{(\mathcal{S}, \mathbf{y})}, \text{ with} \\ \mathcal{R}_{(\mathcal{S}, \mathbf{y})} &= \left\{ (\tilde{\mathbf{w}}, \tilde{\mathbf{r}}, P_{\max}^{\tilde{\mathcal{F}}}) | (\tilde{\mathbf{r}}, \tilde{\mathcal{F}}) \in \mathcal{R}, -\frac{1}{N} \log \Pr \left\{ \mathbf{y} | \mathbf{x}_{(\tilde{\mathbf{w}}, \tilde{\mathbf{r}})}, P_{\max}^{\tilde{\mathcal{F}}} \right\} < \tau_{(\tilde{\mathbf{r}}, P_{\max}^{\tilde{\mathcal{F}}}, \mathcal{S})}(\mathbf{y}) \right\}, \end{aligned} \quad (45)$$

where $\tau_{(\tilde{\mathbf{r}}, P_{\max}^{\tilde{\mathcal{F}}}, \mathcal{S})}(\cdot)$ is the typicality threshold function. Again, we will first analyze the error performance for each individual \mathcal{S} and then derive the overall error performance by taking the union over all \mathcal{S} .

For a given user subset $\mathcal{S} \subset \{1, \dots, K\}$, the following probability terms are defined.

First, assume that (\mathbf{w}, \mathbf{r}) is transmitted over channel $P_{Y|X} \in \mathcal{F}$, with $(\mathbf{r}, \mathcal{F}) \in \mathcal{R}$. Let $P_{t[\mathbf{r}, \mathcal{F}, P_{Y|X}, \mathcal{S}]}$ be the probability that the likelihood value of the transmitted codeword vector calculated using $P_{\min}^{\mathcal{F}}$ is no larger than the corresponding typicality threshold,

$$P_{t[\mathbf{r}, \mathcal{F}, P_{Y|X}, \mathcal{S}]} = \Pr \left\{ P(\mathbf{y} | \mathbf{x}_{(\mathbf{w}, \mathbf{r})}, P_{\min}^{\mathcal{F}}) \leq e^{-N\tau_{(\mathbf{r}, P_{\min}^{\mathcal{F}}, \mathcal{S})}(\mathbf{y})} \right\}. \quad (46)$$

Define $P_{m[(\mathbf{r}, \mathcal{F}), (\tilde{\mathbf{r}}, \tilde{\mathcal{F}}), P_{Y|X}, \mathcal{S}]}$ as the probability that the likelihood value of the transmitted codeword vector calculated using $P_{\min}^{\mathcal{F}}$ is no larger than that of another codeword (\mathbf{w}, \mathbf{r}) with $(\tilde{\mathbf{w}}_{\mathcal{S}}, \tilde{\mathbf{r}}_{\mathcal{S}}) = (\mathbf{w}_{\mathcal{S}}, \mathbf{r}_{\mathcal{S}})$, $(\tilde{w}_k, \tilde{r}_k) \neq (w_k, r_k), \forall k \notin \mathcal{S}$, calculated using $P_{\max}^{\tilde{\mathcal{F}}}$ with $(\tilde{\mathbf{r}}, \tilde{\mathcal{F}}) \in \mathcal{R}$,

$$\begin{aligned} P_{m[(\mathbf{r}, \mathcal{F}), (\tilde{\mathbf{r}}, \tilde{\mathcal{F}}), P_{Y|X}, \mathcal{S}]} &= \Pr \left\{ P(\mathbf{y} | \mathbf{x}_{(\mathbf{w}, \mathbf{r})}, P_{\min}^{\mathcal{F}}) \leq P(\mathbf{y} | \mathbf{x}_{(\tilde{\mathbf{w}}, \tilde{\mathbf{r}})}, P_{\max}^{\tilde{\mathcal{F}}}) \right\} \\ (\tilde{\mathbf{w}}, \tilde{\mathbf{r}}, \tilde{\mathcal{F}}), (\tilde{\mathbf{r}}, \tilde{\mathcal{F}}) &\in \mathcal{R}, (\tilde{\mathbf{w}}_{\mathcal{S}}, \tilde{\mathbf{r}}_{\mathcal{S}}) = (\mathbf{w}_{\mathcal{S}}, \mathbf{r}_{\mathcal{S}}), (\tilde{w}_k, \tilde{r}_k) \neq (w_k, r_k), \forall k \notin \mathcal{S}. \end{aligned} \quad (47)$$

Second, assume that $(\tilde{\mathbf{w}}, \tilde{\mathbf{r}})$ is transmitted over channel $\tilde{P}_{Y|X} \in \tilde{\mathcal{F}}$, with $(\tilde{\mathbf{r}}, \tilde{\mathcal{F}}) \notin \mathcal{R}$. Define $P_{i[(\tilde{\mathbf{r}}, \tilde{\mathcal{F}}), (\mathbf{r}, \mathcal{F}), \tilde{P}_{Y|X}, \mathcal{S}]}$ as the probability that the decoder finds a codeword (\mathbf{w}, \mathbf{r}) with $(\mathbf{w}_{\mathcal{S}}, \mathbf{r}_{\mathcal{S}}) = (\tilde{\mathbf{w}}_{\mathcal{S}}, \tilde{\mathbf{r}}_{\mathcal{S}})$, $(w_k, r_k) \neq (\tilde{w}_k, \tilde{r}_k), \forall k \notin \mathcal{S}$, over channel class \mathcal{F} with $(\mathbf{r}, \mathcal{F}) \in \mathcal{R}$, such that its likelihood value calculated using $P_{\min}^{\mathcal{F}}$ is larger than the corresponding typicality threshold,

$$\begin{aligned} P_{i[(\tilde{\mathbf{r}}, \tilde{\mathcal{F}}), (\mathbf{r}, \mathcal{F}), \tilde{P}_{Y|X}, \mathcal{S}]} &= \Pr \left\{ P(\mathbf{y} | \mathbf{x}_{(\mathbf{w}, \mathbf{r})}, P_{\min}^{\mathcal{F}}) > e^{-N\tau_{(\mathbf{r}, P_{\min}^{\mathcal{F}}, \mathcal{S})}(\mathbf{y})} \right\}, \\ (\mathbf{w}, \mathbf{r}, \mathcal{F}), (\mathbf{r}, \mathcal{F}) &\in \mathcal{R}, (\mathbf{w}_{\mathcal{S}}, \mathbf{r}_{\mathcal{S}}) = (\tilde{\mathbf{w}}_{\mathcal{S}}, \tilde{\mathbf{r}}_{\mathcal{S}}), (w_k, r_k) \neq (\tilde{w}_k, \tilde{r}_k), \forall k \notin \mathcal{S}. \end{aligned} \quad (48)$$

Consequently, the system error probability P_{es} can be upper-bounded using the above probabilities terms as follows,

$$P_{es} \leq \max \left\{ \begin{aligned} &\max_{(\mathbf{r}, P_{Y|X}) : P_{Y|X} \in \mathcal{F}, (\mathbf{r}, \mathcal{F}) \in \mathcal{R}} \sum_{\mathcal{S} \subset \{1, \dots, K\}} \left[P_{t[\mathbf{r}, \mathcal{F}, P_{Y|X}, \mathcal{S}]} \right. \\ &\quad \left. + \sum_{(\tilde{\mathbf{r}}, \tilde{\mathcal{F}}) \in \mathcal{R}, \tilde{\mathbf{r}}_{\mathcal{S}} = \mathbf{r}_{\mathcal{S}}} P_{m[(\mathbf{r}, \mathcal{F}), (\tilde{\mathbf{r}}, \tilde{\mathcal{F}}), P_{Y|X}, \mathcal{S}]} \right], \\ &\max_{(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}) : \tilde{P}_{Y|X} \in \tilde{\mathcal{F}}, (\tilde{\mathbf{r}}, \tilde{\mathcal{F}}) \notin \mathcal{R}} \sum_{\mathcal{S} \subset \{1, \dots, K\}} \left[\sum_{(\mathbf{r}, \mathcal{F}) \in \mathcal{R}, \mathbf{r}_{\mathcal{S}} = \tilde{\mathbf{r}}_{\mathcal{S}}} P_{i[(\tilde{\mathbf{r}}, \tilde{\mathcal{F}}), (\mathbf{r}, \mathcal{F}), \tilde{P}_{Y|X}, \mathcal{S}]} \right] \end{aligned} \right\}. \quad (49)$$

Note that we have used the union bound over all user subsets \mathcal{S} to obtain the probability bound in (49). Next, we will derive individual bound for each of the probability terms on the right hand side of (49).

A derivation similar to (26)-(30) in Appendix A gives the upper bound on $P_{m[(\mathbf{r}, \mathcal{F}), (\tilde{\mathbf{r}}, \tilde{\mathcal{F}}), P_{Y|\mathbf{X}}, \mathcal{S}]}$ as,

$$\begin{aligned} P_{m[(\mathbf{r}, \mathcal{F}), (\tilde{\mathbf{r}}, \tilde{\mathcal{F}}), P_{Y|\mathbf{X}}, \mathcal{S}]} &\leq e^{N\rho \sum_{k \notin \mathcal{S}} \tilde{r}_k} \sum_{\mathbf{y}} E_{\boldsymbol{\theta}_{\mathcal{S}}} \left[E_{\boldsymbol{\theta}_{\tilde{\mathcal{S}}}} \left[P(\mathbf{y}|\mathbf{x}_{(w, \mathbf{r})}, P_{Y|\mathbf{X}}) P(\mathbf{y}|\mathbf{x}_{(w, \mathbf{r})}, P_{\min}^{\mathcal{F}})^{-s} \right] \right. \\ &\quad \times \left. \left[E_{\boldsymbol{\theta}_{\tilde{\mathcal{S}}}} \left[P(\mathbf{y}|\mathbf{x}_{(\tilde{w}, \tilde{\mathbf{r}})}, P_{\max}^{\tilde{\mathcal{F}}})^{\frac{s}{\rho}} \right] \right]^{\rho} \right] \\ &\leq \exp \left\{ -NE_m(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}}, \mathcal{F}, \tilde{\mathcal{F}}) \right\}, \end{aligned} \quad (50)$$

where $E_m(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}}, \mathcal{F}, \tilde{\mathcal{F}})$ is given in (10). Note that the second inequality in (50) is due to the fact that $P(\mathbf{y}|\mathbf{x}_{(w, \mathbf{r})}, P_{Y|\mathbf{X}}) \leq P(\mathbf{y}|\mathbf{x}_{(w, \mathbf{r})}, P_{\max}^{\mathcal{F}})$, and the right hand side of (50) is not a function of $P_{Y|\mathbf{X}}$.

Similarly, by using the same bounding techniques as in (31)-(33) and (34)-(38) in Appendix A, we can upper bound $P_{t[\mathbf{r}, \mathcal{F}, P_{Y|\mathbf{X}}, \mathcal{S}]}$ for any $s_1 > 0$ by,

$$\begin{aligned} P_{t[\mathbf{r}, \mathcal{F}, P_{Y|\mathbf{X}}, \mathcal{S}]} &\leq \sum_{\mathbf{y}} E_{\boldsymbol{\theta}_{\mathcal{S}}} \left[E_{\boldsymbol{\theta}_{\tilde{\mathcal{S}}}} \left[P(\mathbf{y}|\mathbf{x}_{(w, \mathbf{r})}, P_{Y|\mathbf{X}}) P(\mathbf{y}|\mathbf{x}_{(w, \mathbf{r})}, P_{\min}^{\mathcal{F}})^{-s_1} \right] e^{-Ns_1 \tau_{(\mathbf{r}, P_{\min}^{\mathcal{F}}, \mathcal{S})}(\mathbf{y})} \right] \\ &\leq \sum_{\mathbf{y}} E_{\boldsymbol{\theta}_{\mathcal{S}}} \left[E_{\boldsymbol{\theta}_{\tilde{\mathcal{S}}}} \left[P(\mathbf{y}|\mathbf{x}_{(w, \mathbf{r})}, P_{\max}^{\mathcal{F}}) P(\mathbf{y}|\mathbf{x}_{(w, \mathbf{r})}, P_{\min}^{\mathcal{F}})^{-s_1} \right] e^{-Ns_1 \tau_{(\mathbf{r}, P_{\min}^{\mathcal{F}}, \mathcal{S})}(\mathbf{y})} \right], \end{aligned} \quad (51)$$

and upper bound $P_{i[(\tilde{\mathbf{r}}, \tilde{\mathcal{F}}), (\mathbf{r}, \mathcal{F}), \tilde{P}_{Y|\mathbf{X}}, \mathcal{S}]}$ for any $s_2 > 0, 0 < \tilde{\rho} \leq 1$ by,

$$\begin{aligned} P_{i[(\tilde{\mathbf{r}}, \tilde{\mathcal{F}}), (\mathbf{r}, \mathcal{F}), \tilde{P}_{Y|\mathbf{X}}, \mathcal{S}]} &\leq \sum_{\mathbf{y}} E_{\boldsymbol{\theta}_{\mathcal{S}}} \left[E_{\boldsymbol{\theta}_{\tilde{\mathcal{S}}}} \left[P(\mathbf{y}|\mathbf{x}_{(\tilde{w}, \tilde{\mathbf{r}})}, \tilde{P}_{Y|\mathbf{X}}) \right] \right. \\ &\quad \times E_{\boldsymbol{\theta}_{\tilde{\mathcal{S}}}} \left\{ \left[P(\mathbf{y}|\mathbf{x}_{(w, \mathbf{r})}, P_{\min}^{\mathcal{F}})^{\frac{s_2}{\rho}} \right]^{\tilde{\rho}} e^{Ns_2 \tau_{(\mathbf{r}, P_{\min}^{\mathcal{F}}, \mathcal{S})}(\mathbf{y})} e^{N\tilde{\rho} \sum_{k \notin \mathcal{S}} r_k} \right\} \Bigg] \\ &\leq \max_{(\mathbf{r}', \mathcal{F}') \notin \mathcal{R}, \mathbf{r}'_{\mathcal{S}} = \mathbf{r}_{\mathcal{S}}} \sum_{\mathbf{y}} E_{\boldsymbol{\theta}_{\mathcal{S}}} \left[E_{\boldsymbol{\theta}_{\tilde{\mathcal{S}}}} \left[P(\mathbf{y}|\mathbf{x}_{(w', \mathbf{r}'), P_{\max}^{\mathcal{F}'}}) \right] \right. \\ &\quad \times E_{\boldsymbol{\theta}_{\tilde{\mathcal{S}}}} \left\{ \left[P(\mathbf{y}|\mathbf{x}_{(w, \mathbf{r})}, P_{\min}^{\mathcal{F}})^{\frac{s_2}{\rho}} \right]^{\tilde{\rho}} e^{Ns_2 \tau_{(\mathbf{r}, P_{\min}^{\mathcal{F}}, \mathcal{S})}(\mathbf{y})} e^{N\tilde{\rho} \sum_{k \notin \mathcal{S}} r_k} \right\} \Bigg]. \end{aligned} \quad (52)$$

Note that the upper bound given in (51) is not a function of $P_{Y|\mathbf{X}}$. Similarly, the bound in (52) is not a function of $\tilde{P}_{Y|\mathbf{X}}$.

Optimization of the typicality threshold $\tau_{(\mathbf{r}, P_{\min}^{\mathcal{F}}, \mathcal{S})}$ can be carried out using the similar technique as introduced in (39)-(41) in Appendix A. By substituting the optimal $\tau_{(\mathbf{r}, P_{\min}^{\mathcal{F}}, \mathcal{S})}$ into (51) and (52), we get

$$P_{t[\mathbf{r}, \mathcal{F}, P_{Y|\mathbf{X}}, \mathcal{S}]} P_{i[(\tilde{\mathbf{r}}, \tilde{\mathcal{F}}), (\mathbf{r}, \mathcal{F}), \tilde{P}_{Y|\mathbf{X}}, \mathcal{S}]} \leq \max_{(\mathbf{r}', \mathcal{F}') \notin \mathcal{R}, \mathbf{r}'_{\mathcal{S}} = \mathbf{r}_{\mathcal{S}}} \exp \left\{ -NE_i(\mathcal{S}, \mathbf{r}, \mathbf{r}', \mathcal{F}, \mathcal{F}') \right\}, \quad (53)$$

where $E_i(\mathcal{S}, \mathbf{r}, \mathbf{r}', \mathcal{F}, \mathcal{F}')$ is given in (10).

Combining (50), (53) and (49), we obtain

$$P_{es} \leq \max \left\{ \max_{(\mathbf{r}, P_{Y|X}): P_{Y|X} \in \mathcal{F}, (\mathbf{r}, \mathcal{F}) \in \mathcal{R}} \sum_{S \subset \{1, \dots, K\}} \left[\max_{(\mathbf{r}', \mathcal{F}') \notin \mathcal{R}, \mathbf{r}'_S = \mathbf{r}_S} \exp \{ -N E_i(\mathcal{S}, \mathbf{r}, \mathbf{r}', \mathcal{F}, \mathcal{F}') \} \right. \right. \\ \left. \left. + \sum_{(\tilde{\mathbf{r}}, \tilde{\mathcal{F}}) \in \mathcal{R}, \tilde{\mathbf{r}}_S = \mathbf{r}_S} \exp \{ -N E_m(\mathcal{S}, \mathbf{r}, \tilde{\mathbf{r}}, \mathcal{F}, \tilde{\mathcal{F}}) \} \right], \right. \\ \left. \max_{(\tilde{\mathbf{r}}, \tilde{P}_{Y|X}): \tilde{P}_{Y|X} \in \mathcal{F}, (\tilde{\mathbf{r}}, \tilde{\mathcal{F}}) \notin \mathcal{R}} \sum_{S \subset \{1, \dots, K\}} \left[\sum_{(\mathbf{r}, \mathcal{F}) \in \mathcal{R}, \mathbf{r}_S = \tilde{\mathbf{r}}_S} \max_{(\mathbf{r}', \mathcal{F}') \notin \mathcal{R}, \mathbf{r}'_S = \mathbf{r}_S} \exp \{ -N E_i(\mathcal{S}, \mathbf{r}, \mathbf{r}', \mathcal{F}, \mathcal{F}') \} \right] \right\} \quad (54)$$

Since the upper bounds given in (50) and (53) are not functions of individual channels (but functions of channel classes), the right hand side of (54) can be simplified to the right hand side of (9).

REFERENCES

- [1] J. Luo and A. Ephremides, *A New Channel Coding Approach for Random Access with Bursty Traffic*, to appear in IEEE Trans. on Inform. Theory.
- [2] J. Massey and P. Mathys, *The Collision Channel Without Feedback*, IEEE Trans. on Inform. Theory, Vol. IT-31, pp. 192-204, Mar. 1985.
- [3] Z. Wang and J. Luo, *Error Performance of Channel Coding in Random Access Communication*, submitted to IEEE Trans. on Inform. Theory. <http://arxiv.org/abs/1010.0642>.
- [4] A. Lapidoth and P. Narayan, *Reliable Communication under Channel Uncertainty*, IEEE Trans. Inform. Theory, Vol. 44, pp. 2148-177, Oct. 1998.
- [5] D. Blackwell, L. Breiman, and A. Thomasian, *The capacity of A Class of Channels*, Ann. Math. Statist., Vol. 30, pp. 1229-241, 1959.
- [6] J. Wolfowitz, *Simultaneous Channels*, Arch. Rational Mech. and Anal., Vol. 4, pp. 371-86, 1959.
- [7] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd Ed., Wiley Interscience, 2005.
- [8] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, New York: Academic, 1981.
- [9] D. Bertsekas and R. Gallager, *Data Network*, 2nd Ed., Prentice Hall, 1992.
- [10] S. Shamai, I. Teletar, and S. Verdú, *Fountain Capacity*, IEEE Trans. Inform. Theory, Vol. 53, pp. 4372-4376, Nov. 2007.